Phase II Cancer Trials: When and How

NCIC CTG Course for New Investigators
August 9-12, 2011
Learning Objectives

At the end of the session the participant should be able to

• Define the objectives of “screening” vs. “definitive” trials
• Describe the possible endpoints for phase II screening trials
• Understand basic concepts of phase II design including:
  – Non-randomized two-stage designs
  – Ha, Ho, alpha and beta errors in sample size determination
  – Types of randomized phase II design and their possible uses
• Understand the role of correlative studies within phase II screening trials
• Understand some of the controversial aspects of phase II designs for trials of molecular targeted agents.
Phase II Trials: Outline

- Role of phase II trials
- Objectives
- Endpoints
- Patient Population
- Design
- Special considerations:
  - Targeted
  - Randomized phase II trials
  - Correlative studies
Translation: From Laboratory Hypothesis to New Therapy

**PRECLINICAL**

- Laboratory efficacy
- Preclinical toxicology

**CLINICAL**

- Dose determination
- Pharmacokinetics

**TRIAL**

- I
- II
- III

Definitive assessment of efficacy
## Examples: two new agents

### Erlotinib
- **EGFR tyrosine kinase inhibitor**
- **Oral phase I trial:**
  - 150 mg po daily tolerable
  - toxicities: rash, diarrhea
- **Question:**
  - Does it show activity in **ovarian cancer**, a disease with high frequency of EGFR overexpression?

### CCI-779
- **mTOR inhibitor: theoretically of interest when PTEN loss**
- **IV phase I trial:**
  - 25 mg IV weekly tolerable
  - toxicities: rash, mucositis
- **Question:**
  - Does it show activity in **endometrial cancer**, a disease with high frequency of PTEN mutation?
To Demonstrate Efficacy

• **Screening trials** – *does agent merit more study?*
  - Phase II studies
  - “Intermediate” endpoints (e.g. objective response)

• **Definitive trials** – *should this agent be adopted into practice?*
  - Phase III studies
  - Definitive, clinically meaningful endpoints (e.g. survival)
Objectives of Phase II Trials

• Primary:
  – To estimate level of **anti-tumour activity** of an agent or regimen in a given tumour type

• Secondary
  – To provide (further) information on **toxicity**
  – If applicable and possible, to generate hypotheses about relationship of **features of drug target** in tumours and response (or progression).
What are features of optimal **endpoint** for a screening trial?

1. Measures an effect on tumour
2. Standard definition of “effect”
3. Unlikely seen as part of natural history
4. Relatively early event
5. Experience shows *it can reliably identify drugs active in phase III*
Phase II endpoints: Options

1. Objective response (e.g. RECIST)
2. Minor response
3. Proportion non-progressive (non-PD rate)
4. Progression free survival
5. Tumour marker
6. Other biomarker
7. Functional Imaging
8. Some measure of “area under the curve” of maximal % change in tumour size.
Traditional Phase II endpoint: Objective Response

- Using this endpoint in *single agent* trials:
  - Does **not** require randomized design (since tumour shrinkage only rarely spontaneous)
  - Has been reasonably successful in identifying drugs that can improve survival
Standard Response Criteria

Varies by tumour type. Examples:

- **Most Solid Tumours:**
  - Objective response (e.g. RECIST 1.1)

- **Some Solid Tumours:**
  - CA125 response (Ovarian cancer)
  - PSA response (Prostate cancer)
  - MacDonald Criteria (Brain tumours)

- **Hematological malignancies:**
  - Lymphoma
  - IWG AML criteria
New response evaluation criteria in solid tumours: Revised RECIST guideline (version 1.1)

E.A. Eisenhauer\textsuperscript{a,*}, P. Therasse\textsuperscript{b}, J. Bogaerts\textsuperscript{c}, L.H. Schwartz\textsuperscript{d}, D. Sargent\textsuperscript{e}, R. Ford\textsuperscript{f}, J. Dancey\textsuperscript{g}, S. Arbuck\textsuperscript{h}, S. Gwyther\textsuperscript{i}, M. Mooney\textsuperscript{g}, L. Rubinstein\textsuperscript{g}, L. Shankar\textsuperscript{g}, L. Dodd\textsuperscript{g}, R. Kaplan\textsuperscript{j}, D. Lacombe\textsuperscript{c}, J. Verweij\textsuperscript{k}
RECIST 1.1: Measuring Disease

- **Measurable lesion:**
  - $\geq 10 \text{ mm longest diameter on CT scan}$
    (assuming slice thickness 5 mm)
  - $\geq 15 \text{ mm shortest diameter for lymph node}$

- Up to **5** largest measurable lesions assessed (maximum **2** per organ site)

- **Sum of diameters**
RECIST 1.1: Defining Response

- **Complete Response (CR):**
  - Disappearance of all disease
- **Partial Response (PR):**
  - > 30% decrease in sum of diameters
- **Progression (PD):**
  - > 20% increase in sum of diameters **and** at least 5 mm absolute increase
- **Stable disease (SD):**
  - Neither PD nor PR.
- **CR, PR must be confirmed if response primary endpoint.** SD has protocol defined “minimum” duration.
Example of Marker Response: CA125 Response

• Gynecologic Cancer Intergroup criteria*:
  – Patient must have one baseline elevated sample (at least 2x ULN)
  – CA125 response if:
    • 50% fall from baseline
    • Confirmed by repeat sample at least 28 days later

Patient Population

- Patients to be enrolled in phase II trial should have characteristics which:
  - Allow assessment of primary endpoint in patients with disease of interest
  - Maximize the chance of seeing activity
  - Take into account drug toxicity and pharmacology
Examples: Population

CCI-779: Endometrium

- Measurable disease
- Performance status (ECOG) 0,1,2.
- No prior chemotherapy. One hormonal treatment allowed.

Erlotinib: Ovary

- Measurable disease; +/- CA125 > 2x ULN
- Able to swallow; no bowel obstruction
- One prior chemo regimen. Two cohorts will be studied:
  - > 6 mo
  - < 6 mo
Design

- Design should do two things:
  - Allow identification of truly active drug (i.e. limit the risk of a false negative result)
  - Limit the number of patients treated in case the drug is truly inactive.
“Classic” Single Arm Design

- Multistage (usually 2-stage) non-randomized study.

- Sample size and stopping rule based on the level of activity (response rate) of interest (\(H_a\)) and the levels of the 2 key error rates:
  - The \(\alpha\) error: false positive result
  - The \(\beta\) error: false negative result — mostly we want to minimize this.
Statistical Design/Sample Size

• Several methods available:
  – Simon, Fleming, Gehan….

• Consult with statistician

• In general:

  Smaller:
  Response rate (Ha)
  $\alpha$ value
  $\beta$ value

  Larger:
  Sample Size
Examples: Design

**CCI-779: Endometrium**
- **Ha**: 20%
- **Ho**: 5%
- Enter 15 patients
  - Close trial if no responses
  - If $\geq 1$ response: enroll 15 additional pts
- If $\geq 4/30$ pts respond conclude agent is of interest for further study

$\alpha = 0.058; 1 - \beta = 0.87$

**Erlotinib: Pt. Sens. Ovary**
- **Ha**: 30%
- **Ho**: 5%
- Enter 8 patients
  - Close trial if no responses
  - If $\geq 1$ response: enroll 7 additional pts
- If $\geq 3/15$ pts respond conclude agent is of interest for further study

$\alpha = 0.03; 1 - \beta = 0.85$

*Platinum resistant: as at left*
Reporting Results

- Account for all patients entered

- Describe:
  - Patient characteristics
  - Treatment delivery
  - Toxic effects
  - No. pts with: CR, PR, SD, PD
  - Response rate: based on all eligible patients (do not inflate response rate by reducing denominator)
  - Response Duration
  - Outcome of any “special” endpoint e.g. molecular marker
After Phase II is Complete:

• If a minimum level of activity seen further evaluation warranted:
  – Confirmatory phase II
  – Combination phase I/II – may be randomized
  – Randomized single agent studies

• If no responses, drug concluded to be of no interest for further study
Issues with “Traditional” Phase II

• What about…

  -- Targeted anti-cancer drugs (so called non-cytotoxic agents) that may not cause tumour shrinkage in animals?
Era of Molecular Targeted Therapy: Phase II Trial Challenges

- Is response a realistic endpoint with molecular targeted agents?

- If not:
  - What are options of alternative endpoints?
  - What implications do use of non-response endpoints have on design?

- Should population be “enriched” for target expression?
Phase II endpoints:

Options

1. Objective response (e.g. RECIST)
2. Minor response
3. Proportion non-progressive (non-PD rate)
4. Progression free survival
5. Tumour marker
6. Other biomarker
7. Functional Imaging
8. Some measure of “area under the curve” of maximal % change in tumour size.
SU11248 Maximum % Reduction of Target Lesions by Patient

- Partial Responders by RECIST
- SD/PD Patients
Phase II endpoints: Options

1. Objective response (e.g. RECIST)
2. Minor response
3. Proportion non-progressive (non-PD rate)
4. Progression free survival
5. Tumour marker
6. Other biomarker
7. Functional Imaging
8. Some measure of “area under the curve” of maximal % change in tumour size.

Require randomized designs
What alternative endpoints/designs have been used in Phase II screening trials of targeted agents?

- 19 targeted agents with single agent phase II reports in one of 6 common tumour types
- Describe endpoints, design used and whether results related to eventual phase III success
- Total: 89 trials found
## Agents/Targets: Phase II Review

<table>
<thead>
<tr>
<th>Target</th>
<th>Agent</th>
<th>#Reports</th>
</tr>
</thead>
<tbody>
<tr>
<td>Angiogenesis</td>
<td>ZD6474</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>SU5416</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Bevacizumab</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>SU11248</td>
<td>2</td>
</tr>
<tr>
<td>PKC alpha</td>
<td>ISIS 3521</td>
<td>5</td>
</tr>
<tr>
<td>raf kinase</td>
<td>BAY 43-9006</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>ISIS 5132</td>
<td>4</td>
</tr>
<tr>
<td>DNA MTase</td>
<td>MG98</td>
<td>1</td>
</tr>
<tr>
<td>MEK</td>
<td>CI-1040</td>
<td>1</td>
</tr>
<tr>
<td>mTOR</td>
<td>CCI-779</td>
<td>3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Target</th>
<th>Agent</th>
<th>#Reports</th>
</tr>
</thead>
<tbody>
<tr>
<td>EGFR/HER2</td>
<td>ZD1839</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>OSI-774</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>C225</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>trastuzumab</td>
<td>8</td>
</tr>
<tr>
<td>ckit/abl</td>
<td>STI571</td>
<td>7</td>
</tr>
<tr>
<td>MMP</td>
<td>Marimastat</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>BMS-275291</td>
<td>1</td>
</tr>
<tr>
<td>Farnesyl trans-</td>
<td>R115777</td>
<td>5</td>
</tr>
<tr>
<td>ferase</td>
<td>SCH66336</td>
<td>1</td>
</tr>
</tbody>
</table>
## Results: Study Outcomes by Trial

<table>
<thead>
<tr>
<th>Tumour Type</th>
<th>Not reported</th>
<th>0</th>
<th>&gt;0 - ≤10</th>
<th>&gt;10-≤20</th>
<th>&gt;20</th>
</tr>
</thead>
<tbody>
<tr>
<td>Breast</td>
<td>0</td>
<td>6</td>
<td>5</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>Colorectal</td>
<td>0</td>
<td>8</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Lung</td>
<td>2</td>
<td>10</td>
<td>4</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Ovary</td>
<td>0</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Prostate</td>
<td>7</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Renal</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td><strong>TOTAL (%)</strong></td>
<td><strong>13 (15)</strong></td>
<td><strong>38 (43)</strong></td>
<td><strong>19 (21)</strong></td>
<td><strong>11 (12)</strong></td>
<td><strong>8 (9)</strong></td>
</tr>
</tbody>
</table>

Note: The objective response rate is calculated as the proportion of trials reporting objective responses.
### Summary and Comparison of Cytotoxics with Targeted Agents Reviewed

<table>
<thead>
<tr>
<th>Response rates (all trials)</th>
<th>No. Targeted Agents</th>
<th>No. Approved for ANY Tumour type Included in Review %</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt;20</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>10 - &lt;20</td>
<td>2 (3)</td>
<td>67</td>
</tr>
<tr>
<td>&gt;0 - &lt;10</td>
<td>6</td>
<td>67</td>
</tr>
<tr>
<td>TOTAL</td>
<td>19</td>
<td>7 (8)</td>
</tr>
</tbody>
</table>

\[ p = 0.005 \text{ (excluding gefitinib)} \]

\[ p < 0.0001 \text{ (including gefitinib)} \]
Randomized Phase II Designs

• Small sample size trials which are not adequately powered to compare outcomes.

• Endpoints can be response or other (e.g. PFS) depending on design and question being addressed

• Useful in several circumstances....
Examples: Randomized Phase II Designs

- "Pick the Winner" Two schedules of a new drug look interesting. Randomized trial to select the schedule most likely to be best:

  e.g. IND.163

Patients:
Breast cancer
1 prior chemotherapy

RAD001 daily
RAD001 weekly
Examples: Randomized Phase II Designs

- **Identify early evidence of effect:** Standard vs. experimental regimen to identify sufficient activity to merit phase III.

  e.g. BR.20

  Patients: SCLC, Completed Rx with CR, PR

  Primary endpoint: 2.5 mo improvement in PFS

  Statistics: power 80%, one-sided alpha 0.1
Examples: Randomized Phase II Designs

- Use control arm to interpret results:
  Control arm serves to help *interpretation of results* in experimental arm
  
  e.g. IND.165

Patients:
CRPC
No prior chemotherapy

Primary endpoint: PSA response rate

Statistics:
  - $H_0 < 40\%, H_1 > 60\%, \alpha = 0.1, \beta = 0.1$
  - $>20/40$ pts in OGX-011 Arm with PSA resp of interest
Correlative Biology/Enrichment

• When to enrich for molecularly defined subset?

• What to enrich for?

• Depends on how robust the preclinical data is on molecular predictor

• In all cases: need tumour collection from all patients
An Easy Example

- Agent designed to inhibit mutated variant of target → select for patients having tumours with mutated target.

Vemurafenib (PLX 4032)
More Common Setting

- Agent affects *normal variant of target(s)*
- Most drugs affect several targets
- Phase II trial: opportunity to develop/refine hypotheses about which tumour subsets are most sensitive (or least insensitive)
- May be challenging – look for altered expression, mutations, amplifications in pathway and in salvage pathway
**Example:**

Phase II Single Agent Trial mTOR inhibitor in Endometrial ca

- PTEN IHC positive = normal
- Negative = loss of PTEN expression
IND.160 A
(Temsieolimus Endometrium)

<table>
<thead>
<tr>
<th>Response</th>
<th>PTEN +ve</th>
<th>PTEN -ve</th>
</tr>
</thead>
<tbody>
<tr>
<td>PR</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>SD</td>
<td>3</td>
<td>14</td>
</tr>
<tr>
<td>PD</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>IN</td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>Total</td>
<td>9</td>
<td>20</td>
</tr>
</tbody>
</table>

No evidence PTEN loss needed for PR
Similar results pmTOR, pS6k

PTEN IHC positive = normal
Negative = loss of PTEN expression
# Mutational Analysis

<table>
<thead>
<tr>
<th>Best Response Association Tested</th>
<th>Mutation type</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>PR vs. other</td>
<td>Any</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td>PI K3CA</td>
<td>0.4</td>
</tr>
<tr>
<td></td>
<td>PI K3CA or Akt</td>
<td>1.0</td>
</tr>
<tr>
<td>PR + SD vs. other</td>
<td>Any</td>
<td>0.3</td>
</tr>
<tr>
<td></td>
<td>PI K3CA</td>
<td>0.4</td>
</tr>
<tr>
<td></td>
<td>PI K3CA or Akt</td>
<td>0.7</td>
</tr>
<tr>
<td>PD vs. other</td>
<td>Any</td>
<td>0.7</td>
</tr>
<tr>
<td></td>
<td>PI K3CA</td>
<td>0.7</td>
</tr>
<tr>
<td></td>
<td>PI K3CA or Akt</td>
<td>0.7</td>
</tr>
</tbody>
</table>
Lesson: It is challenging to find “THE” biomarker

• Phase II can be place to start but seldom where the story finishes

• Sample size is relatively small to do much more than generate/refine hypotheses
Summary: Phase II Trials

- **Goal**: Screen new agent/combination for activity.
- **Primary Endpoints**:
  - Single agent/single arm: Objective response
  - Randomized (single agent or combo): response or PFS
- **Design**:
  - Depends if single arm or randomized
  - Should minimize possibility of false negative outcome
- **Correlative studies**: Should be routine
  - At minimum: to generate hypotheses about selection biomarker
- **Drugs active in phase II**: Need further evaluation